

R. BRUCE BRIGGS

Manager of Programming and Operations
Campus Computing Network
University of California
Los Angeles

The User Interface For Bibliographic Search Services

The user interface for a bibliographic search service includes the entire set of user-service interactions, from initial phrasing of the information needs to final review of the results. This interaction has occurred in reference libraries for many years, and an investigation of the literature in this area yields a wealth of information. Analysis of this information, however, shows that while the functions of the reference process are fairly well defined, the process itself is not, and is essentially left to the judgment of the reference librarian.

When analysis is limited to literature on computer-based bibliographic search services, a similar picture emerges. This literature deals primarily with technical and operational problems, as might be expected, because these were the initial problems facing the designers and developers of this comparatively new service. Attention was given to the user interface, but not with the same studied approach that was given to other components of the system. Nevertheless, as solutions to the technical and operational problems are being found, the importance of understanding the user interface is being increasingly appreciated. The desire to serve the user better, the attention given to automating components of the user interface, and the growing interest in information networks of libraries and bibliographic search services are just some of the reasons for attaining this understanding.

In December 1973, the University of Georgia (UGA) and UCLA began a

joint research project with funding from the National Science Foundation to study the user interface for computer-based bibliographic search services. There were two objectives of the study: (1) to study the existing user interfaces and to model the search services at UGA and UCLA after them, and (2) to propose one or more models for the user interface for a multidisciplinary bibliographic information network. The study teams at each institution were composed of staff who had participated in the design and development of the respective bibliographic search centers.

The search center at UGA is known as the Georgia Information Dissemination Center (GIDC). It serves twenty-five institutions of higher education in the University System of Georgia, as well as numerous governmental agencies, other academic institutions, and commercial organizations throughout the United States and in several other countries. The search center at UCLA is known as the Center for Information Services (CIS), and serves twenty-two institutions of higher education in the University of California and California State Universities and Colleges system, as well as numerous private colleges and universities, governmental agencies, and commercial organizations throughout the United States. The GIDC operation is the larger of the two, offering current awareness searches on eleven data bases and retrospective searches on sixteen data bases; CIS presently offers only current awareness searches on five data bases.

Both centers provide trained reference librarians or information specialists to assist the user in obtaining service from the search center. GIDC has four information specialists on its staff who handle search queries primarily for the University of Georgia, Georgia State University, commercial users, and the smaller two- and four-year institutions in Georgia. These four staff members have graduate degrees in the subject matter areas appropriate to their data base specializations; none has had previous library work experience. Search profiles for users at the Georgia Institute of Technology and the Medical College of Georgia, as well as for users at remote sites in Ohio and New York, are prepared by reference librarians at those locations who have received workshop training on computer-based retrieval from the GIDC staff. CIS has two staff members whose primary responsibilities are to train reference librarians in the various institutions served, and to advise the reference librarians how to handle difficult or unusual search requests. These two staff members occasionally work directly with users who do not otherwise have access to a trained specialist; however, the vast majority of users work through the forty-five reference librarians in the institutions that CIS serves.

RESEARCH STUDY METHODOLOGY

The approach used for this research study was for the teams at each institution to operate independently, but in parallel. The purpose was to

achieve a certain degree of redundancy, to create a check-and-balance function, and to provide a means for determining intercenter differences. The two teams maintained a close interaction by meeting regularly during the period of the study.

The study was divided into three phases: data collection, analysis, and model development. The data collection phase was designed to collect information on all aspects of the user interface, with an emphasis on the interaction process between the user and the reference librarian or information specialist (hereafter called the intermediary). All data collection instruments used were developed for purposes of this study; some were developed jointly and used by both teams (to provide a basis for intercenter comparisons), while others were developed and used by one of the two teams. All instruments were field-tested prior to the beginning of the data collection period. The jointly developed data collection instruments were:

User Information Form This form was completed by users prior to the user-intermediary interview, and included information such as: a prose formulation of the user's question; suggested keywords, synonyms, and excluded terms; search type requested (current awareness or retrospective); and information about the user (e.g., previous experience with computer-based search systems, occupation, intended use of results, etc.).

Post-Interview Questionnaire for Intermediaries This form was completed by the intermediary immediately following an interview with a user. Its purpose was to gather information on the intermediary's perception of different aspects of the interview, such as the user's information needs, the nature of the user's question, the user's attitude, the expected performance of the profile, and the general nature of the interview.

Follow-Up Questionnaire for Users This form was sent to each user involved in the study following receipt of at least one set of search results from each data base specified for the user's profile. The questionnaire was intended to obtain the user's general reactions to the search service, to the usefulness of the results, and to the elapsed time from the interview to receipt of results. Users were also asked to make additional comments and suggestions.

Both teams also tape-recorded user-intermediary interviews and transcribed them for subsequent analysis. Users' permission was obtained beforehand, and each was assured that the tapes would be used solely for research purposes.

The analysis phase began about two-thirds of the way through the data collection period with a preliminary analysis of data collected at that time to test procedures. The analysis was, in part, data-directed, and had the objectives of describing in detail the components of the user interface, and of providing the basis for development of the models. The particular analysis

techniques used were independently selected by the two study teams and then compared for completeness and consistency.

It was not known at the outset whether the analysis would lead to the specification of one or more than one model of the existing user interface. Conceivably, there could be a different model for each center's interface, or even a different model for each intermediary. In any event, after modeling the existing user interface, the objective was then to propose one or more models for the user interface for an information network.

RESULTS

The complete results of this study are presently being compiled for the final report to the National Science Foundation and were not available at the time this paper was written. Therefore, the results presented in this paper, unless explicitly stated otherwise, are based on the UCLA study team's findings. These results are grouped into six major categories: user characteristics, intermediary characteristics, interview process, users' questions, profile coding, and search results.

User Characteristics

The user population is predominantly academic with 44 percent of the users being graduate students, 37 percent academic researchers, and 33 percent faculty members. (The categories overlap in some cases because users were asked to check all applicable titles.) The nonacademic segment is primarily staff (17 percent of all users), which consists of administrators and technical and general support staff.

Results regarding the purpose of the search also reflect the predominantly academic character of UCLA's users: 74 percent of the profile searches were performed for research projects, 36 percent for master's degree theses, 27 percent for teaching purposes, 27 percent for bibliographies for publications, and 32 percent for personal bibliographies. Less popular uses of the search results were seminars (10 percent), class projects (11 percent), term papers (8 percent), doctoral dissertations (7 percent), and patent searches (2 percent). Eighty-eight percent of the users were new to this particular retrieval system. Most of them (95 percent) had heard of computer-based retrieval, but 56 percent indicated no previous experience with it and only about one-tenth (11 percent) were highly experienced.

The users' attitudes and expectations were relatively good: most (71 percent) exhibited an interested (optimistic) attitude toward computer-based retrieval; the majority (88 percent) expected the service to be of some value, but not perfect. Most (85 percent) of the users did not change their attitude during the interview.

The most significant difference reflected by the data collected was the user's estimated ability to write his own profile. While 79 percent of the users believed they could write their own profiles, the intermediaries thought only 48 percent of the users could do so, and the researchers thought only 36 percent could do so. However, more than one-half of the users who believed they could write their own profiles qualified it by saying this was true if they had help from the intermediary. Fifty-four percent of the users indicated they would want to write their own profiles. It is interesting to note the lack of any significant correlation between users estimating that they could write their own profiles and wanting to write their own profiles. Seventy-seven percent of the users responding said they could write their own profiles (on their own, with an analyst, or with a good manual), but 33 percent said that although they could do it, they did not want to.

Intermediary Characteristics

Of the thirty-eight intermediaries who responded to this questionnaire, 81 percent have graduate degrees in library science and 74 percent are currently working as reference librarians. All seven of the intermediaries who do not have library science degrees do have graduate degrees in other fields and/or experience with library systems and data processing.

The intermediaries' profile writing experience, including experience with other computerized information retrieval systems, ranged from less than three months to more than two years. More than one-half of the intermediaries have been writing profiles for over one year, with 34 percent having more than two years of experience. From a list of five alternatives (apprenticeship, workshops or seminars, self-taught, manuals, and other), the intermediaries were asked to check all applicable means by which they were trained for writing profiles. The responses indicated that the two most frequent means were through workshops or seminars (76 percent) and profile writing manuals (71 percent). Sixty-three percent said they were self-taught and 29 percent received training through an apprenticeship. Fifteen percent listed other means of training, e.g., library school courses.

The intermediaries identified five major problem areas which specifically related to their roles as intermediaries between users and CIS:

1. *Users* Seven of the intermediaries listed problems getting the users to participate, specifically in expressing their needs in the initial interview, in providing additional information for inclusion in the profile, and in providing feedback concerning satisfaction with the citations retrieved.
2. *CIS* Seven intermediaries listed as a problem the delays and slow turn-

around time from the time the profile is submitted to CIS to the arrival of the first set of citations.

3. *Time* Ten intermediaries reported that the added time needed for user interviews, profile construction and coding, output review, and profile revisions presented a problem since they were also expected to continue to perform their regular professional duties as reference librarians. Mention was also made of the burden of the additional time which had to be devoted to "clerical" tasks such as recordkeeping, forwarding output to users, etc.
4. *Training* Seven intermediaries listed the need for more training by CIS staff in profile construction and profile refinement, with emphasis on the individual data bases. Intermediaries also felt that they needed more technical information on how the computerized information retrieval system works to better prepare them to answer the broad range of questions asked by the users.
5. *Subject Knowledge* Although many intermediaries referred to the problem throughout the questionnaire, five intermediaries specifically mentioned that the main problem they encountered was that the required depth and breadth of the subject knowledge to write successful profiles was much greater than that needed to be a "successful" reference librarian.

In light of the problems listed by the intermediaries, it was surprising that a significant controversy appeared in their responses to questions which sought their opinions on automating parts of the user interview and profile construction process. There was a 50-50 split (15 percent did not answer this question) between the intermediaries who thought parts of it could and should be automated and those who felt (some rather strongly) that no part of the process should be automated. None of the intermediaries felt that the entire interview could or should be automated; however, many suggested supplementary on-line capabilities, i.e., thesauri, sample data bases and syntax checking, which could be used during the interview and profile construction and coding process.

The Interview Process

EVENT ANALYSIS

Transcripts from eight interviews tape-recorded for test purposes were reviewed to produce a list of significant events that occur in the user-inter-

<i>Event Number</i>	<i>Event</i>	<i>Number of Occurrences</i>
25	INFORMATION REQUEST CLARIFIED BY USER	259
24	ANALYST REQUESTS CLARIFICATION OF INFORMATION NEEDS	236
38	SEARCH STRATEGY DISCUSSED	232
55 U	USER REQUESTS TUTORIAL	160
28	SEARCH TERM SELECTED BY USER (PROMPTED)	143
18 A	DISCUSS SUBJECT AREA TERMINOLOGY AND RETRIEVAL CAPABILITY	126
27	SEARCH TERM SELECTED BY USER (VOLUNTEERED)	79
59	INDEXING POLICY OF DATA BASE DESCRIBED	100
67	OTHER	99
18 B	USER TUTORIAL ON SUBJECT	90
37	VOCABULARY (PROFILE CODING AIDS) CONSULTED	80
17	REFER TO USER INFORMATION FORM (OTHER THAN REQUEST STATEMENT)	78
6	SEARCH PROCEDURE DESCRIBED	77
5 A	SEARCH SYSTEM DESCRIBED	62
29	SEARCH TERM SELECTED BY ANALYST (USER AGREEMENT)	52
26	SEARCH TERM SELECTED FROM REQUEST FORM	37

Figure 1. Most Frequent Events

mediary interview. This analysis produced a list of seventy-two "expected" events. Then, with appropriate checks to minimize inconsistencies due to subjective judgments by the study team, each transcript was analyzed and the identified events were recorded in the margin in their chronological sequence. These event sequences were then analyzed to determine whether patterns within the sequences existed.

Computer programs were written to yield frequency listings of individual events, and of subsequences (see Figure 1). The analysis then proceeded to determine the relationships between events, e.g.: How often does event A precede event B, and how often does it follow event B? A surprising result of this analysis was that if event B had the highest frequency of following event A, and event C had the highest frequency of following event B, the frequency of the sequence A, B, C was not similarly high. This observation was consistent for all event sequences.

The study team then tried to produce a graph where the nodes were the events, and the arcs indicated the number of times (above a given threshold) that one event preceded or followed another. The result was dubbed the "spaghetti model" by the UCLA team (see Figure 2). Examining the matrix associated with this graph, the team found two salient characteristics: (1) the matrix is not sparse, and (2) most values are small. These characteristics are indicative of a nondeterministic process.

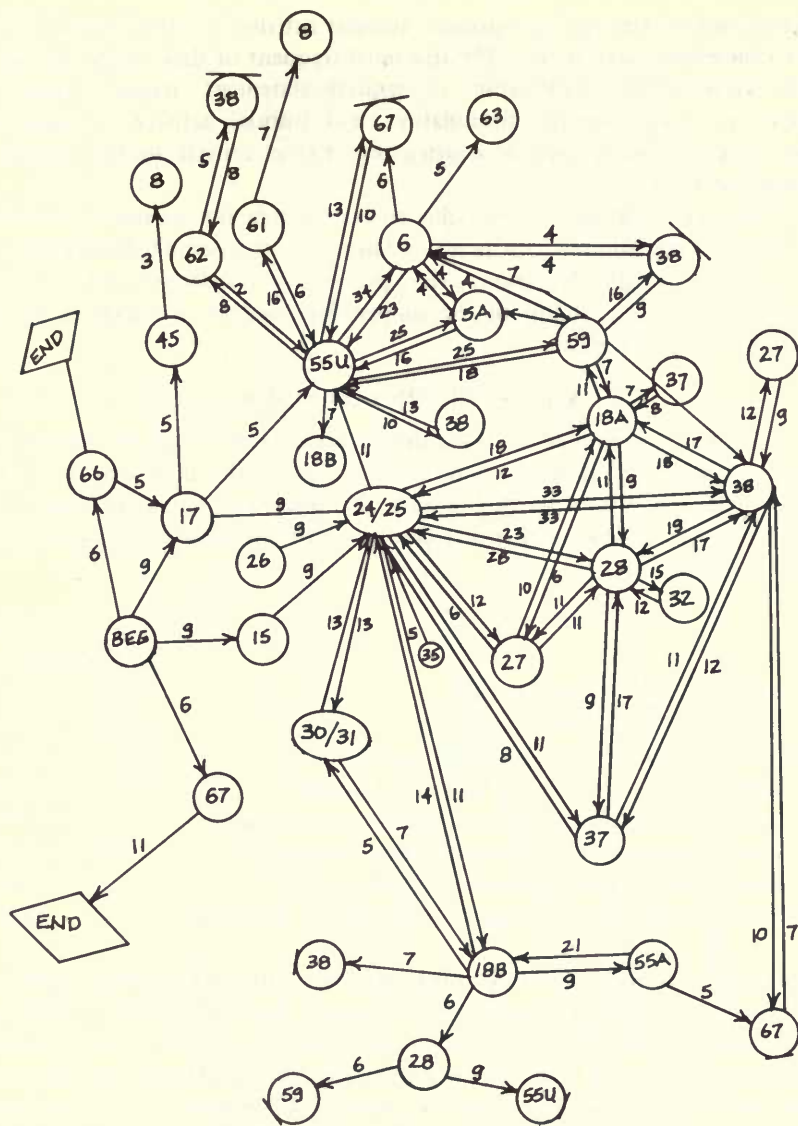


Figure 2. Graph of Event Sequences

Using a different approach, the team grouped events into categories which were descriptive of the functions of the events, resulting in the following categories: system description, data base selection, search type selection, clarification of request statement, request statement negotiation, profile con-

struction, search strategy formulation, tutorial activity, diverting activity, and other (miscellaneous) activity. The five most frequent of these categories were profile construction, clarification of request statement, request statement negotiation, search strategy formulation, and tutorial activity. A graph of these categories shows profile construction to be central to the interview process (see Figure 3).

The conclusion drawn from this analysis is that the interaction between the user and the intermediary is not a linear process, despite descriptions to the contrary in the literature. The process is nondeterministic, and is characterized by the ability of the human intermediary to adapt it to the particular needs of the user.

VALUE OF THE INTERVIEW

The study attempted to assess the value of the interview by determining whether it was helpful to the intermediary and to the user, and by determining what conditions existed during the interview which influenced its helpfulness. The following elements correlated to interviews which were helpful to the intermediary:

	Kendall Coefficient
User was prepared for the interview	.77041
User expressed his literature needs well	.65731
User left interview with an optimistic attitude	.48618
User actively participated in the interview process	.47495
User understood and participated in profile construction	.43951
User was knowledgeable in his subject area literature	.42780
User had written a clear information request	.38563
User was important in term suggestion during the interview	.26558

Interestingly, the data shows that, although many of the users asked questions and provided information freely when discussing their subject area, clarifying their search requests, etc., they become less active and needed to be prompted more when the activity centered on profile construction. During the interview, other than during profile construction activity, 45 percent of the users were curious and willing to cooperate with the intermediary and asked some questions, but gave information mainly when prompted. Only 11 percent of the users seemed reluctant to ask questions and gave short answers to the intermediaries' questions. During the profile construction activity, only 15 percent of the users were interested in the details of profile construction, understood Boolean logic and citation term scanning, and were able to participate in the profile construction process at this level. Fifty percent of

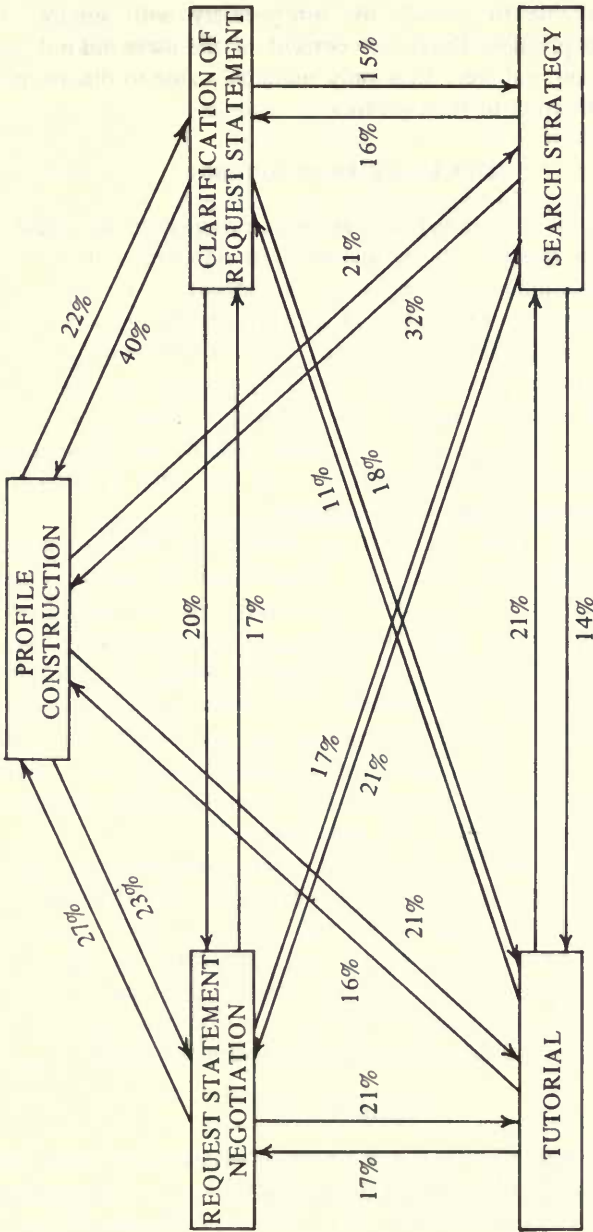


Figure 3. Interaction Between Major Event Categories

the users had a lesser understanding of profile retrieval mechanisms and, although they were able to discuss and participate on a less sophisticated level, they were able to provide the intermediary with suitable terms for inclusion in their profiles. Thirty-five percent of the users did not understand profile construction and were thus only minimally able to discuss and participate in the construction of their profiles.

The Users' Questionnaire

On the user information form, users were requested to prepare a prose statement of their question. In evaluating these statements, the intermediaries perceived no discernible weaknesses in 49 percent of the cases, while the researchers perceived no weaknesses in 41 percent. The perceived weaknesses in the remaining cases were predominantly that the statements were either too broad (18 percent) or too narrow (27 percent). In very few cases did the user either not know what he wanted (8 percent), or expect too much of the retrieval service (6 percent). The researchers judged that 81 percent of the written statements of interests were clear, but that only 40 percent of them were complete enough for profile preparation.

The scope of users' questions as observed by the intermediaries and researchers was split into broad (37 percent and 42 percent respectively), average (33 percent and 29 percent, respectively) and narrow (28 percent and 29 percent, respectively). When forced to choose, users preferred a broad search (89 percent) for high recall to a narrow search (11 percent) emphasizing precision. Both the intermediaries (76 percent) and the researchers (69 percent) confirm this in their perception of user preferences. (The user's stated preference and the intermediary's observation of it correlate with a coefficient of .31 and significance of .001; researchers' perceptions correlate with users' with a coefficient of .38, significance .001.)

Although UCLA is not presently offering a retrospective search service, 17 percent of the users specified that they wanted only this type of service; 25 percent specified that they wanted only current awareness service; and 58 percent specified both. Intermediaries and researchers tended to categorize most users (79 percent and 59 percent, respectively) as wanting either a moderate number of references (50-100) or a large number (greater than 100) and did not view users' expectations as being substantially different.

Profile Coding

Nearly two-thirds (63 percent) of the profiles were coded completely after the interview, and 27 percent of the profiles were coded partially during the interview and completed after the interview. Therefore, a total of 90

percent of the profiles were coded partly or entirely after the interview. Typically, then, when the intermediaries begin to code a profile, they have received input from the user in writing via the user information form and in person via the interview. The intermediaries indicated they may also consult other resources for help in coding the profile.

The study team attempted to ascertain the extent to which all the identified potential inputs to the profile are included in the profile. They began by looking at the coded profile and by diagramming its structure. Then, they became familiar with the user information form and the interview transcript, and described the profile's correlation to both these forms of user input.

The profile and the user information form are:

<i>Rating</i>	<i>Percentage</i>
very similar	15
generally similar	31
slightly similar	23
dissimilar	29
unable to determine because of technical nature of subject	0
not reported	2

The profile correlates to:

<i>Rating</i>	<i>Percentage</i>
the user information form almost exclusively	4
the user information form more than the interview	13
both about equally	27
the interview more than the statement of interest	36
the interview almost exclusively	2
neither very much	17

Thus, in the study team's judgment, the profile is generally similar to the user information form less than one-half (46 percent) of the time, and only slightly similar or dissimilar in about one-half (52 percent) of the cases studied. Also, while 76 percent of the profiles are derived from both the user information form and the interview, the interview is more important in a larger number of cases (38 percent as compared to 17 percent). Sixty-five percent of the profiles were described as having been derived strictly (13 percent) or largely (52 percent) from user-supplied information. Twenty-seven percent were derived largely from the intermediary's ideas but generally confirmed by the user in the interview. Eight percent were based primarily on the intermediary's concepts and apparently not confirmed by the user. Regarding term selection, the user contributes either primarily (50 percent) or in cooperation with the intermediary (40 percent) in 90 percent of the profiles; the intermediary

contributes either primarily (10 percent) or with the user (40 percent) in 50 percent of the profiles.

Besides the intermediary and the user, what other inputs to the profile coding process are there? Word counts show that 32 percent of a profile's subject terms are not specifically mentioned either in the interview transcript or on the user information form. These search terms may be derived from the intermediary's previous knowledge of the subject, from various printed resources, or from conversations with other persons. The data concerning these sources of input to the profile are somewhat scattered, but can be summarized as follows:

1. For 51 percent of the profiles, the intermediaries describe their previous knowledge of the subject as generally familiar (41 percent) or very familiar (10 percent).
2. For 80 percent of the profiles, the intermediaries report using profile coding aids (thesauri, word frequency lists, coding manuals).
3. Texts or reference books were consulted by the intermediary while coding 36 percent of the profiles.
4. Dictionaries were consulted in order to code 37 percent of the profiles.
5. Personal contacts, besides the user, were resources employed in 13 percent of the profiles.
6. Other resources were consulted for 18 percent of the profiles.
7. During the interview, hard copies of the available data base(s) were consulted in 21 percent of the cases; other profiling aids were used during 28 percent of the interviews.

It is also relevant to note that the percentage of profile words from sources other than the user information form and the interview apparently does not affect the profile success, i.e., there is no statistical correlation.

There is a large degree of variation in the time spent by the intermediary in constructing the profile. The mean is about 1½ hours, but the median is just a little over one hour. The range is from ten minutes to seven hours, with 75 percent of the profiles taking between one-half hour and two hours.

The 44 profiles studied in detail search the following data bases: Current Index to Journals in Education (38 percent), Research in Education (38 percent), Social Science Citation Index (38 percent), Biological Abstracts (67 percent), BioResearch Index (58 percent), Chemical Abstracts—organic chemistry sections (38 percent), Chemical Abstracts—inorganic chemistry sections (15 percent), and CAIN—agricultural citations (4 percent).

Sixty-nine percent of the profiles search tapes produced by different suppliers. However, only 48 percent are constructed to search differently on

one or more data bases; 23 percent use different primary retrieval mechanisms; 13 percent use different subject vocabularies.

The primary retrieval mechanisms used in the coded profiles are: subject words in citation titles and indexing phrases (98 percent), subject codes (44 percent), authors (60 percent), referenced citations (21 percent), journal titles or abbreviations (15 percent), and subject words in citation abstracts (6 percent). The profiles include the following refinements (some are not possible on certain data bases): language restrictions (23 percent), journal names—included (13 percent), journal names—excluded (4 percent), and publication type—patent, book, articles, etc. (2 percent).

After the profiles were coded, the intermediaries expected that 40 percent of them would yield approximately the number of references the user wanted, 31 percent would yield more, and 25 percent would yield less. They also expected that 46 percent would need revisions for improved precision, 28 percent would need improved recall, 22 percent would need no revisions, and 1 percent would need to be completely rewritten.

Search Results

The information obtained on search results will be presented in three sections: user reaction, intermediary reaction, and feedback and revisions.

USER REACTION

The overall reaction of the users to their search results was fairly positive in many respects:

1. Ninety-one percent of the users described their search result as either very useful (37 percent) or of some use (54 percent).
2. Eighty-one percent reported that references were retrieved which were not previously known.
3. Fifty-six percent indicated that they considered the retrieval system a fast means of performing the search; 69 percent indicated that the search had saved them time.
4. Although 39 percent of the users received fewer relevant citations than expected, 64 percent reported the number of citations was about right to be manageable and useful.
5. Eighty-nine percent of the users ranked convenience among the three most important of the things they liked. Thoroughness and the printed compilation of references were also ranked important by more than one-half the users.

Significant negative reactions by the users included the following:

1. About one-half (48 percent) of the users described the elapsed time from construction of the profile to receipt of the answers as too slow, but only one user said it was not useful due to slowness. Thirty-three ranked slowness among their top three dislikes.
2. A total of 31 percent of the users were dissatisfied with the number of citations; 19 percent said they received too few and 12 percent reported too many.
3. Forty-four percent of the profiles could, in the users' opinions, benefit from the revisions; however, only 25 percent wanted revisions made soon; others said they did not have time, it was too late, or it was not worth the effort.
4. Twenty-five percent of the users were not sure if they will continue the search service; 8 percent said they do not expect to continue.
5. The most frequently ranked dislike was "no way to judge completeness." It was ranked among the top three important dislikes by 65 percent of the users (41 percent ranked it most important).

There are two indications that users are more severely discouraged by too few references than they are by too many. All of the users reporting too many answers still described the search as of some use. But 59 percent of the users reporting too few or no answers found their result of little or no use. Nearly all users reporting too many answers indicated that revisions were in order, but about one-half of the users with too few or no hits felt they did not have time to determine needed revisions, or it was not worth the effort, or it was too late to be of help to them. Understandably, users getting too many citations hold more hope for getting useful results from their profile than users receiving too few or no hits. The correlations between questions on the user follow-up survey reveal an interesting phenomenon: a user tends typically to be either generally pleased with the search results and system or generally disillusioned and critical.

INTERMEDIARY REACTION

Results from the first several searches are usually mailed to the intermediary for review purposes before being forwarded to the user. The intermediary reviews the output, notes problems and needed revisions, and then forwards the output to the user. Records kept by intermediaries participating in this study show that they spend an average of forty minutes in reviewing the output for each profile; thus, the intermediary usually has an idea of the usefulness of the search output independent of any feedback received from the user.

The intermediaries were fairly positive in their evaluation of the search results, although slightly less so than the users themselves. Twenty-one percent

of the profiles were rated very successful by the intermediaries, 43 percent moderately successful, 19 percent marginally successful, and 4 percent unsuccessful. The intermediaries felt the success of 12 percent of the profiles was as yet undetermined when the questionnaire was collected.

Intermediaries indicated that 55 percent of the profiles usually retrieved references which were satisfactory both in number and relevance. They indicated that 14 percent retrieved many irrelevant references, 22 percent retrieved few relevant references, and 4 percent retrieved no references. It was interesting to observe that the intermediaries, like the users, react more unfavorably to too few hits than to too many. Fifty-three percent of the profiles receiving "few relevant hits" were described at best as only marginally successful, but 73 percent of the profiles receiving "many irrelevant hits" were described at worst as "moderately successful."

FEEDBACK AND REVISIONS

A considerable amount of activity after the first search results is well within the norm for CIS profiles. There is some indication that even more activity—e.g., output review, discussion with users, and profile revision—would result in more satisfied users. There is also indication that more participation by the users may be beneficial.

Intermediaries spent an average total of 41 minutes reviewing output for each profile; the range was 0 to 185 minutes. Note that the data are somewhat inexact, since intermediaries had received different quantities of search results at the time the data were collected. During the data collection period, 39 percent of the profiles were revised. For the revised profiles, an average of forty-eight minutes was spent by the intermediaries making from one to eight textual revisions per profile. The predominant number of revisions was one; however, the extent of the changes included in each revision submitted can vary greatly. Only 6 percent of the profiles were completely rewritten.

The need for revisions does not correlate with overall usefulness; 34 percent of the profiles studied were described as useful but with revisions still needed. However, there is a relationship between revisions needed and the number of answers received; when users described their searches as receiving approximately the right number of answers to be useful they tend not to require revisions as frequently as those receiving too many or too few. Thus, a significant number of users are finding their results useful, but are desiring revisions to bring the number of citations down or up to the desired level.

For most of the revised profiles, the user suggested changes or worked with the intermediary on the revisions. User participation was apparently a motivation for revisions. The researchers found postsearch activity by the user notably lacking, describing 44 percent of the profiles as "needs to be reviewed by the intermediary and user together." This apparent lack of communication

regarding the search result is further evidenced by: (1) the reported lack of feedback from users to intermediaries (intermediaries cited lack of user feedback as affecting the success of 34 percent of the profiles; the researchers observed that only 33 percent of the users gave feedback to their intermediary); (2) for 39 percent of the profiles, the intermediaries reported that the need for revisions was undetermined at the time of the survey; (3) for 12 percent of the profiles, the intermediaries reported the profile success as undetermined; and (4) for 9 percent of the profiles, the intermediaries reported the quality and quantity of hits retrieved as unknown. The inability of the intermediary to assess the need for revisions and the lack of user feedback correlate quite strongly; about one-half of the profiles described as affected by the lack of user feedback could not be assessed by the intermediary in terms of needed revisions. Similarly, about one-half of the profiles for which the intermediary could not determine the need for revisions were affected by the lack of user feedback.

MODEL OF THE EXISTING USER INTERFACE

The model of the existing user interface was developed jointly by UGA and UCLA, and is based on the analysis of the data collected in the study. The model, as shown in Figure 4, can be considered to have two major components: the presearch activities, represented by the upper half of the figure; and the postsearch activities, represented by the lower half. Both of these components involve interactive processes and, based on the findings of the study, are shown as nondeterministic processes.

In the presearch activities, administrative procedures generally occur first, and profile entry and editing occur last. No predictive statements can be made about the order of the remaining activities, except to say that they are generally interspersed throughout the presearch activities. The study showed that both the informational activities and the data base and search type selection activities were largely supportive of other activities.

The study did not collect and analyze as much data on the postsearch activities as it did on the presearch activities; however, the data collected did indicate that this portion of the user interface was also nondeterministic. The supportive activities occur in this component also because of the tendency of users to add or drop data bases, to request a current awareness search after a retrospective one (or vice versa), and to require additional information as to the need for revisions.

This model describes the existing user interfaces at both UGA and UCLA, and may be descriptive of the interface at other search centers. The interpretation of the model will, of course, vary from center to center, and even from case to case within a center.

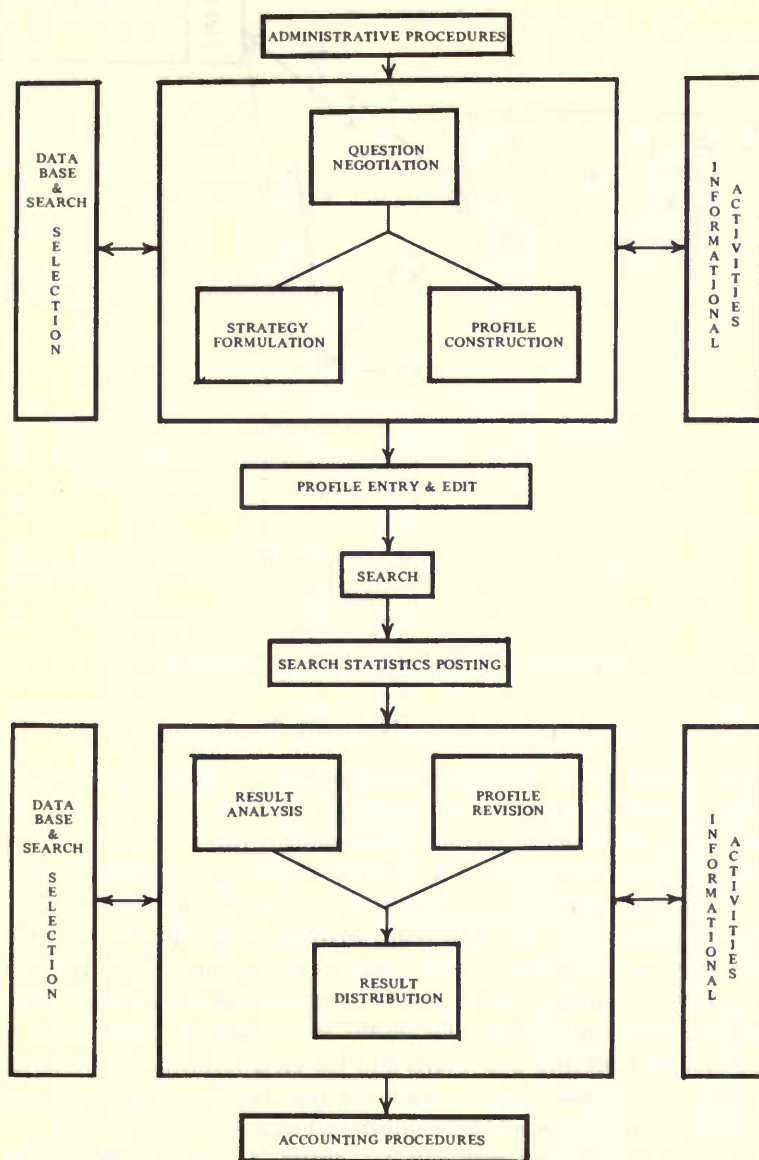


Figure 4. Model of Existing User Interface

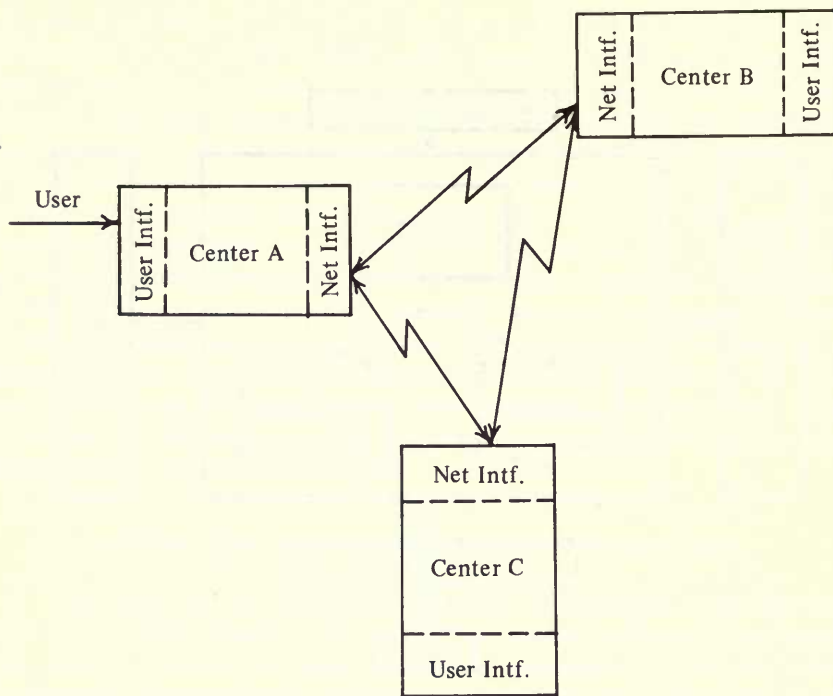


Figure 5. Proposed Network Configuration

User Interface for a Network Environment

The study teams considered a number of different network configurations before choosing to propose the one shown in Figure 5. In this configuration each center in the network maintains two interfaces, one for the user and one for the network. The purpose of this user interface is the same as for existing user interfaces—to transform the user's query into an effective profile, and to assist the user with result analysis and profile revisions. In fact, this interface can be similar to the center's existing one. The network interface provides a standard mechanism for communications among centers by transforming outgoing profiles and results into exchange formats, by transforming incoming profiles and results into the local formats, and by serving as a mechanism for remote users to reach the user interface. The structure of a search center in this environment is shown in Figure 6.

For example, a user may contact center A and request searches of several data bases, some of which are available at center A and the remainder of which are available at center B. The user's query is transformed into a profile in center A's format as a result of the user's interaction with center

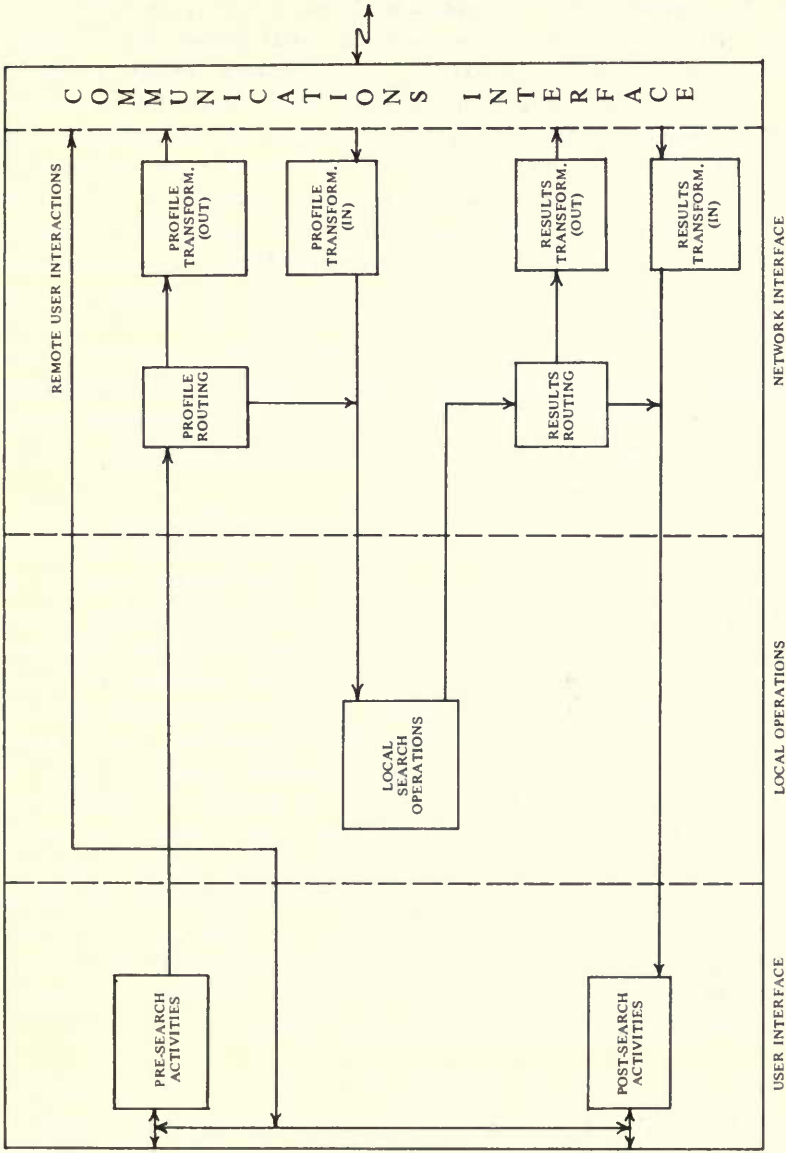


Figure 6. Search Center in Network Environment

A's user interface. This profile is then used to search against those data bases available at center A. Concurrently, the profile is transformed by center A's network interface to an exchange format and sent to center B for searching against the remaining data base specified. Upon receipt, center B's network interface transforms the profile into center B's local format and processes it. Results from center B are transformed to an exchange format and sent to center A where they are transformed to the local format for printing and delivery to the user. There are several advantages to this approach: users interact with only one interface—that of the center most convenient to them—but still have access to all of the services available in the network; centers can maintain their existing user interface; the configuration is modular, allowing for expansion or contraction of the network; and the configuration is adaptable—changes to one center's user interface do not necessitate changes throughout the network.

Components of the network identified by this study as requiring standardization prior to the realization of the network are: data base elements and content, the exchange language for profiles, the communications protocol for control parameters (e.g., data bases to be searched, search type, etc.), and the specifications for transmission of results.

One conclusion of this study is that the intermediary serves a necessary and vital role in the user interface. Some users do not, however, have trained intermediaries available to them. In addition, intermediaries indicate that sometimes they need assistance in areas such as profile construction, language features, and processing procedures. For these reasons, a system specialist is proposed for the network environment. The system specialist would have in-depth knowledge of the profile language, profiling aids and data bases, and would have general subject knowledge. Where appropriate, several system specialists with more specialized subject knowledge could be used. The function of the system specialist would be to serve as a consultant and resource person for both users and intermediaries, and to serve as the primary intermediary for users (e.g., remote users) not having access to an intermediary. The method of communication between the specialist and the user could be through interactive terminal messages, via telephone or, as is being experimented with by one existing center, via headphones to users at interactive terminals. The interaction between the user or the primary intermediary and the system specialist can be represented by Figure 7.

This study really only scratches the surface of the user interface problem. The study teams believe that there are a large number of areas that warrant further study, and an equal number of areas that are as yet unstudied. For example, the recorded interviews contain a wealth of information about the interactive process; only a small portion of this information was gleaned

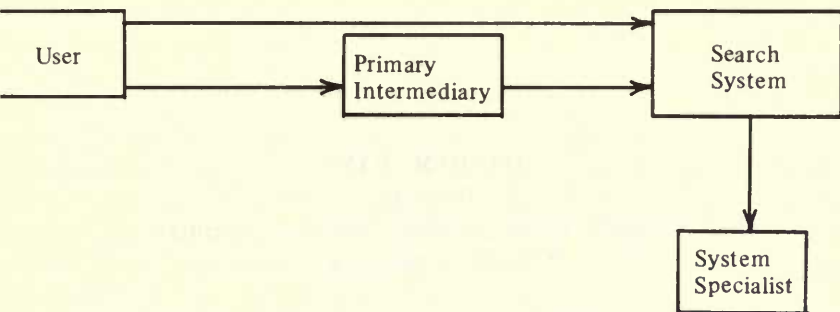


Figure 7. Role of System Specialist

For analysis. Also, the emphasis of this study was on the presearch activities, but the postsearch activities are equally important. Now that the technical and operational problems initially experienced by search centers are under control, attention needs to be turned to the human engineering aspects of the service.